



Patent
Attorney Docket No. 030681-610

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of

Jung-hoe Kim et al.

Group Art Unit: 2819

Application No.: 10/735,894

Examiner: Unassigned

Filing Date: December 16, 2003

Confirmation No.: 5705

Title: METHOD AND APPARATUS FOR ENCODING/DECODING AUDIO DATA WITH SCALABILITY

SUBMISSION OF CERTIFIED COPY OF PRIORITY DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

The benefit of the filing date of the following priority foreign application(s) in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

Country: Republic of Korea

Patent Application No(s).: 10-2002-008032

Filed: December 16, 2002

In support of this claim, enclosed is a certified copy(ies) of said foreign application(s). Said prior foreign application(s) is referred to in the oath or declaration and/or the Application Data Sheet. Acknowledgment of receipt of the certified copy(ies) is requested.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

By



Charles F. Wieland III

Registration No. 33,096

P.O. Box 1404
Alexandria, Virginia 22313-1404
(703) 836-6620

Date: May 24, 2004



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출 원 번 호 : 10-2002-0080320
Application Number

출 원 년 월 일 : 2002년 12월 16일
Date of Application DEC 16, 2002

20396

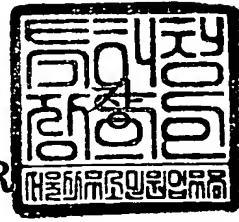
출 원 인 : 삼성전자주식회사
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003 년 12 월 02 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0007
【제출일자】	2002. 12. 16
【국제특허분류】	G06F
【발명의 명칭】	비트율 조절가능한 오디오 부호화 방법, 복호화 방법, 부호화 장치 및 복호화 장치
【발명의 영문명칭】	Method and apparatus for encoding/decoding audio data with scalability
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	1999-009556-9
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2000-002816-9
【발명자】	
【성명의 국문표기】	김중회
【성명의 영문표기】	KIM, Jung Hoe
【주민등록번호】	760206-1231712
【우편번호】	157-014
【주소】	서울특별시 강서구 화곡4동 504-117
【국적】	KR
【발명자】	
【성명의 국문표기】	김상욱
【성명의 영문표기】	KIM, Sang Wook
【주민등록번호】	670228-1009911

【우편번호】 137-042
【주소】 서울특별시 서초구 반포2동 반포주공아파트 202동 201호
【국적】 KR
【발명자】
【성명의 국문표기】 오은미
【성명의 영문표기】 OH, Eun Mi
【주민등록번호】 680318-2051113
【우편번호】 137-044
【주소】 서울특별시 서초구 반포4동 미도2차아파트 503-305
【국적】 KR
【취지】 특허법 제42조의 규정에 의하여 위와 같이 출원합니다. 대리인
이영필 (인) 대리인
이해영 (인)
【수수료】
【기본출원료】 20 면 29,000 원
【가산출원료】 21 면 21,000 원
【우선권주장료】 0 건 0 원
【심사청구료】 0 항 0 원
【합계】 50,000 원
【첨부서류】 1. 요약서·명세서(도면)_1통

【요약서】

【요약】

비트율 조절가능한 오디오 부호화 방법, 복호화 방법, 부호화 장치 및 복호화 장치가 개시된다.

본 발명에 따라 오디오 데이터를 비트율 조절이 가능하도록 부호화하는 방법은 오디오 데이터를 비트율 조절이 가능하도록 부호화하는 방법에 있어서, (a) 복수개의 계층에 상응하도록 오디오 데이터를 분할하는 단계; (b) 상기 복수개의 계층에 각각 상응하는 스케일 밴드 정보 및 코딩 밴드 정보를 얻는 단계; (c) 첫번째 계층에 상응하는 스케일 밴드 정보 및 코딩 밴드 정보에 따른 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 부호화하는 단계; (d) 상기 첫번째 계층에 해당하는 오디오 데이터를 대응하는 스케일 팩터 정보를 참조하여 양자화하여 양자화 샘플을 얻는 단계; (f) 얻어진 복수개의 양자화 샘플을 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 헤프만 부호화하는 단계; 및 (g) 상기 복수개의 계층에 대한 부호화가 완료될 때까지 매 회 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 (c) 단계 내지 (f) 단계를 반복하여 계층 구조를 갖도록 패킹하는 단계를 포함하는 것을 특징으로 한다. 이에 의해, FGS(Fine Grain Scalability)를 제공하면서도 complexity가 보다 낮아진다.

【대표도】

도 7

【명세서】

【발명의 명칭】

비트율 조절가능한 오디오 부호화 방법, 복호화 방법, 부호화 장치 및 복호화 장치{Method and apparatus for encoding/decoding audio data with scalability}

【도면의 간단한 설명】

도 1은 본 발명의 바람직한 실시예에 따른 부호화 장치의 블럭도,

도 2는 본 발명의 바람직한 실시예에 따른 복호화 장치의 블럭도,

도 3은 본 발명에 따라 비트율을 조절할 수 있도록 계층 구조로 부호화된 비트스트림을 구성하는 프레임의 구조도,

도 4는 부가 정보의 상세 구조도,

도 5는 본 발명에 따른 부호화 방식을 개략적으로 설명하기 위한 참고도,

도 6은 본 발명에 따른 부호화 방식을 보다 구체적으로 설명하기 위한 참고도,

도 7은 본 발명의 바람직한 실시예에 따른 부호화 방법을 설명하기 위한 플로우챠트,

도 8은 본 발명의 바람직한 실시예에 따른 복호화 방법을 설명하기 위한 플로우챠트,

도 9는 본 발명의 바람직한 실시예에 따른 복호화 방법을 설명하기 위한 플로우챠트이다

【발명의 상세한 설명】**【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

- <10> 본 발명은 오디오 데이터의 부호화 및 복호화에 관한 것으로, 보다 상세하게는 오디오 데이터를 비트율 조절가능하도록 부호화하는 방법, 복호화하는 방법, 그 부호화 장치 및 복호화 장치에 관한 것이다.
- <11> 최근 디지털 신호처리 기술의 발달에 의해 오디오 신호는 디지털 데이터로 저장되고 재생되는 경우가 대부분이다. 디지털 오디오 저장/재생 장치는 아날로그 오디오 신호를 샘플링하고 양자화하여 디지털 신호인 PCM(Pulse Code Modulation) 오디오 데이터로 변환하여 CD, DVD와 같은 정보저장매체에 저장해둔 다음 사용자가 필요로 할 때 이를 재생해서 들을 수 있도록 해준다. 디지털 방식에 의한 오디오 신호의 저장/복원 방식은 LP(Long-Play Record), 마그네틱 테이프와 같은 아날로그 저장/복원 방식에 비해 음질을 크게 향상시켰고 저장 기간에 따른 열화 현상을 현저히 감소시켰으나 디지털 데이터의 크기가 적지 않아 저장 및 전송이 원활하지 못한 문제점이 있었다.
- <12> 이와 같은 문제점을 해결하기 위해, 디지털 오디오 신호의 크기를 줄이기 위한 다양한 압축 방식이 사용되고 있다. ISO (International Standard Organization)에 의해 표준화 작업이 이루어진 MPEG/audio(Moving Pictures Expert Group)나 Dolby사에 의해 개발된 AC-2/AC-3는 인간의 심리음향 모델(Psychoacoustic Model)을 이용하여 데이터의 양을 줄이는 방법을 채용하였고 그 결과 신호의 특성에 관계없이 효율적으로 데이터의 양을 줄일 수 있었다. 즉,

MPEG/audio 표준이나 AC-2/AC-3 방식은 이전의 디지털 부호화 방식에 비해 1/6 내지 1/8로 줄어든 64 Kbps - 384 Kbps 비트율만으로 CD의 음질과 거의 같은 정도의 음질을 제공한다.

<13> 그러나, 이를 방법은 모두 고정된 비트율에 대해 최적의 상태를 찾아 양자화 과정과 부호화 과정을 거치는 방식을 따르므로, 네트워크를 통해 전송할 때 네트워크 상황이 좋지 않아 전송 대역폭이 낮아지면 끊김이 발생하며 사용자에게 더 이상의 서비스를 제공할 수 없게 되는 문제점이 있다. 또한, 제한된 저장 용량을 가지고 있는 이동식 기기에 적합하도록 좀 더 작은 크기의 비트스트림으로 변환하고자 할 때 크기를 줄이기 위해서는 재부호화 과정을 거쳐야 하므로 많은 계산량이 요구된다.

<14> 이에, 본 출원인은 비트 분할 산술 부호화(BSAC, Bit-Sliced Arithmetic Coding) 기법을 사용하여 비트율 조절이 가능한 오디오 부호화/복호화 방법 및 장치를 1997년 11월 19일자 대한민국 특허출원 제97-61298호로 출원하여 2000년 4월 17일자 등록특허 제261253호로 등록받았다. BSAC에 따르면, 높은 비트율로 부호화된 비트스트림을 낮은 비트율의 비트스트림으로 만들 수도 있고 그 중 일부의 비트스트림만을 가지고도 복원이 가능하므로 네트워크에 과부화가 걸리거나 복호화기의 성능이 좋지 않거나 또는 사용자가 낮은 비트율을 요구하면 비트스트림의 일부만을 가지고도 - 비트율이 낮아진 만큼 성능의 열화를 보이겠지만 - 사용자에게 어느 정도의 음질로 서비스를 제공할 수 있다.

<15> 그러나 BSAC는 산술 부호화(arithmetic coding)를 채용하고 있음으로 인해 압축 개인은 높지만 많은 계산량을 요구하므로 complexity가 높아 실제로 장치에 구현할 때 비용이 증가하는 단점이 있다.

【발명이 이루고자 하는 기술적 과제】

<16> 따라서, 본 발명의 목적은 FGS(Fine Grain Scalability)를 제공하면서도 complexity가 보다 낮은 비트율 조절가능한 오디오 부호화 방법, 복호화 방법, 그 부호화 장치 및 복호화 장치를 제공하는 것이다.

【발명의 구성 및 작용】

<17> 상기 목적은 본 발명에 따라, (a) 복수개의 계층에 상응하도록 오디오 데이터를 분할하는 단계; (b) 상기 복수개의 계층에 각각 상응하는 스케일 밴드 정보 및 코딩 밴드 정보를 얻는 단계; (c) 첫번째 계층에 상응하는 스케일 밴드 정보 및 코딩 밴드 정보에 따른 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 부호화하는 단계; (d) 상기 첫번째 계층에 해당하는 오디오 데이터를 대응하는 스케일 팩터 정보를 참조하여 양자화하여 양자화 샘플을 얻는 단계; (e) 얻어진 복수개의 양자화 샘플을 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 허프만 부호화하는 단계; 및 (f) 상기 복수개의 계층에 대한 부호화가 완료될 때까지 매 회 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 (c)단계 내지 (f)단계를 반복하여 계층 구조를 갖도록 패킹하는 단계를 포함하는 것을 특징으로 하는 부호화 방법에 의해 달성된다.

<18> 상기 (f)단계 이전에 (e) 상기 복수개의 계층에 대해 각각 사용가능한 비트 범위를 구하는 단계를 포함하고, 상기 (f)단계는 부호화된 비트 수를 카운트하여, 카운트된 비트 수가 대응하는 비트 범위를 초과하면 부호화를 중지하여, 할당된 양자화 샘플을 전부 부호화한 이후에도 카운트된 비트 수가 대응하는 비트 범위보다 작으면 하위 계층에서 부호화하지 못하고 남은 비트를 상기 비트 범위가 허용하는 만큼 부호화하는 단계를 포함하는 것이 바람직하다.

<19> 상기 (a)단계는 (a1) 오디오 데이터를 MDCT 변환하는 단계; 및 (a2) MDCT 변환된 오디오 데이터를 컷-오프 주파수를 참조하여 상기 복수개의 계층에 상응하도록 분할하는 단계를 포함하는 것이 바람직하다.

<20> 상기 (f)단계는 (f1) 복수개의 양자화 샘플을 비트 플래인 상에 매핑하는 단계; 및 (f2) 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위만큼 심벌 단위로 부호화하는 단계를 포함하는 것이 바람직하고, 상기 (f1)단계는 N 개의 양자화 샘플을 비트 플래인 상에 매핑하는 단계이고, 상기 (f2)단계는 N 비트의 이진 데이터로 구성된 상기 심벌에 대응하는 하나의 스칼라 값을 구하는 단계; 및 상기 N 비트의 이진 데이터, 구해진 스칼라 값을 상기 비트 플래인 상에서 현재의 심벌에 대한 상위 심벌에 대응하는 스칼라 값을 참조하여 헤프만 부호화를 수행하는 단계를 포함하는 것이 바람직하다(N은 정수).

<21> 한편, 본 발명의 다른 분야에 따르면, 상기 목적은 계층 구조로 부호화된 오디오 데이터를 비트율 조절이 가능하도록 복호화하는 방법에 있어서, (a) 첫번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 복호화하는 단계; (b) 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 헤프만 복호화하여 양자화 샘플을 얻는 단계; (c) 얻어진 양자화 샘플을 상기 스케일 팩터 정보를 참조하여 역양자화하는 단계; (d) 역양자화된 샘플을 MDCT 역변환하는 단계; 및 (e) 미리 결정된 계층까지 복호화가 완료될 때까지 매 희 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 (a)단계 내지 (d)단계를 반복 수행하는 단계를 포함하는 것을 특징으로 하는 복호화 방법에 의해서도 달성된다.

- <22> 상기 (b)단계는 (b1) 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위 만큼 심벌 단위로 복호화하는 단계; 및 (b2) 복호화된 심벌이 배열된 비트 플래인으로부터 양자화된 샘플을 얻는 단계를 포함하는 것이 바람직하다.
- <23> 상기 (b1)단계는 복호화된 심벌로 구성된 $4*N$ 비트 플래인을 얻는 단계이고, 상기 (b2) 단계는 상기 $4*N$ 비트 플래인으로부터 N 개의 양자화 샘플을 얻는 단계를 포함하는 것이 바람직하다(N 은 정수).
- <24> 한편, 본 발명의 다른 분야에 따르면, 상기 목적은 계층 구조로 부호화된 오디오 데이터를 비트율 조절이 가능하도록 복호화하는 장치에 있어서, 첫번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 복호화하고, 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 헤프만 복호화하여 양자화 샘플을 얻는 언패킹부: 얻어진 양자화 샘플을 상기 스케일 팩터 정보를 참조하여 역양자화하는 역양자화부; 및 역양자화된 샘플을 역변환하는 역변환부를 포함하는 것을 특징으로 하는 복호화 장치에 의해서도 달성된다.
- <25> 상기 언패킹부는 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위 만큼 심벌 단위로 복호화하고, 복호화된 심벌이 배열된 비트 플래인으로부터 양자화된 샘플을 얻는 것이 바람직하다.
- <26> 상기 언패킹부는 복호화된 심벌로 구성된 $4*N$ 비트 플래인을 얻은 다음 상기 $4*N$ 비트 플래인으로부터 N 개의 양자화 샘플을 얻는 것이 바람직하다(N 은 정수).

- <27> 한편, 본 발명의 다른 분야에 따르면, 오디오 데이터를 비트율 조절이 가능하도록 부호화하는 장치에 있어서, 상기 오디오 데이터를 MDCT 변환하는 변환부; 각 계층에 해당하는, MDCT 변환된 오디오 데이터를 상기 스케일 팩터 정보를 참조하여 양자화하여 양자화 샘플을 출력하는 양자화부; 및 각 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 부호화하고, 상기 양자화부로부터의 복수개의 양자화 샘플을 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 헤프만 부호화하는 패킹부를 포함하는 것을 특징으로 하는 부호화 장치에 의해서 도 달성된다.
- <28> 상기 패킹부는 복수개의 계층에 각각 상응하는 스케일 밴드 정보 및 코딩 밴드 정보를 얻어 각 계층에 상응하는 스케일 밴드 정보 및 코딩 밴드 정보에 따른 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 부호화하는 것이 바람직하다.
- <29> 상기 패킹부는 부호화된 비트 수를 카운트하여, 카운트된 비트 수가 각 계층에 허용되는 비트 범위를 초과하면 부호화를 중지하여, 할당된 양자화 샘플을 전부 부호화한 이후에도 카운트된 비트 수가 대응하는 비트 범위보다 작으면 하위 계층에서 부호화하지 못하고 남은 비트를 상기 비트 범위가 허용하는 만큼 부호화하는 것이 바람직하다.
- <30> 상기 패킹부는 MDCT 변환된 오디오 데이터를 상기 컷-오프 주파수를 참조하여 복수개의 계층에 상응하도록 분할하는 것이 바람직하고, 복수개의 양자화 샘플을 비트 플래인 상에 매핑하고, 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위만큼 심벌 단위로 부호화하는 것이 바람직하며, K 개의 양자화 샘플을 비트 플래인 상에 매핑하고, K 비트의 이진 데이터로 구성된 상기 심벌에 대응하는 하나의 스칼라 값을 구한 다음 상기 K 비트의 이진 데이터, 구해진 스칼라 값과 상기 비트 플래인 상에

서 현재의 심벌에 대한 상위 심벌에 대응하는 스칼라 값을 참조하여 허프만 부호화하는 것이 바람직하다(K는 정수).

- <31> 이하 첨부된 도면을 참조하여 본 발명의 바람직한 실시예에 대해 상세히 설명한다.
- <32> 도 1은 본 발명의 바람직한 실시예에 따른 부호화 장치의 블럭도이다.
- <33> 도 1을 참조하면, 부호화 장치는 본 발명에 따라 비트율 조절가능하도록 오디오 데이터를 계층 구조로 부호화하는 장치로서, 변환부(11), 심리음향부(12), 양자화부(13) 및 비트패킹부(14)를 포함한다.
- <34> 변환부(11)는 시간 영역의 오디오 신호인 PCM(Pulse Coded Modulation) 오디오 데이터를 입력받아 심리음향부(12)로부터의 제공되는 음향심리모델에 관한 정보를 참조하여 주파수 영역의 신호로 변환한다. 시간 영역에서는 인간이 인지하는 오디오 신호의 특성의 차이가 그리 크지 않지만, 변환을 통해 얻어진 주파수 영역의 오디오 신호는 인간의 음향심리모델에 따라 각 주파수 대역에서 인간이 느낄 수 있는 신호와 느낄 수 없는 신호의 특성 차이가 크기 때문에 각 주파수 대역 별로 할당되는 비트수를 다르게 함으로써 압축의 효율을 높일 수 있다. 본 실시예에서 변환부(11)는 MDCT(Modified Discrete Cosine transform) 변환을 수행한다.
- <35> 심리음향부(12)는 어택(attack) 감지 정보, 등 음향심리모델에 관한 정보를 변환부(11)로 제공하는 한편, 변환부(11)에 의해 변환된 오디오 신호를 적절한 서브 밴드의 신호들로 묶고 각 신호들의 상호작용으로 인해 발생되는 마스킹현상을 이용하여 각 서브 밴드에서의 마스킹 문턱치(masking threshold)를 계산하여 양자화부(13)로 제공한다. 마스킹 문턱치란 오디오 신호들의 상호 작용으로 인해 인간이 들어도 느끼지 못하는 신호의 최대 크기를 말한다. 본

실시예에서 심리음향부(12)는 BMLD(binaural masking level depression)를 이용하여 스테레오 성분에 대한 마스킹 문턱치 등을 계산한다.

<36> 양자화부(13)는 인간이 들어도 느끼지 못하도록 각 대역의 양자화 잡음의 크기가 심리음향부(12)에서 제공된 마스킹 문턱치보다 작도록 각 대역의 오디오 신호들을 대응하는 스케일 팩터 정보를 기초로 스칼라 양자화하여 양자화 샘플들을 출력한다. 즉, 양자화부(13)는 음향 심리부(12)에서 계산된 마스킹 문턱치와 각 대역에서 발생하는 잡음(noise)의 비율인 NMR (Noise-to-Mask Ratio)를 이용하여 전 대역의 NMR 값이 0 dB 이하가 되도록 양자화한다. NMR 값이 0 dB 이하라는 것은 양자화 잡음을 인간이 들을 수 없음을 의미한다.

<37> 비트 패킹부(14)는 각 계층에 속하는 양자화 샘플들 및 부가 정보를 부호화하여 계층 구조로 패킹한다. 부가 정보는 각 계층에 해당하는 스케일 밴드 정보, 코딩 밴드 정보, 그 스케일 팩터 정보 및 코딩 모델 정보를 포함한다. 스케일 밴드 정보와 코딩 밴드 정보는 헤더 정보로서 패킹되어 복호화 장치로 전송될 수도 있고, 각 계층마다의 부가 정보로서 부호화되고 패킹되어 복호화 장치로 전송될 수도 있으며, 복호화 장치에 미리 저장되어 있음으로 인해 전송되지 않을 수도 있다.

<38> 보다 구체적으로, 비트 패킹부(14)는 첫 번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 부호화하는 한편, 첫 번째 계층에 상응하는 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 부호화한다. 다음으로 두 번째 계층에 대해서도 동일한 과정을 반복한다. 즉, 미리 결정된 복수개의 계층에 대한 부호화가 완료될 때까지 계층을 증가시키면서 부호화한다. 본 실시예에서 비트 패킹부(14)는 스케일 팩터 정보와 코딩 모델 정보는 차분 부호화하고, 양

자화 샘플은 허프만 부호화한다. 본 발명에 따라 부호화된 비트스트림의 계층 구조는 후술한다.

<39> 스케일 밴드 정보는 오디오 신호의 주파수 특성에 따라 보다 적절하게 양자화를 수행하기 위한 정보로, 주파수 영역을 복수개의 밴드로 나누고 각 밴드에 적합한 스케일 팩터를 할당하였을 때 각 계층에 대응하는 스케일 밴드를 알려주는 정보를 말한다. 이에, 각 계층은 적어도 하나의 스케일 밴드에 속하게 된다. 각 스케일 밴드는 할당된 하나의 스케일 팩터를 가진다. 코딩 밴드 정보 또한 오디오 신호의 주파수 특성에 따라 보다 적절하게 부호화를 수행하기 위한 정보로, 주파수 영역을 복수개의 밴드로 나누고 각 밴드에 적합한 코딩 모델을 할당하였을 때 각 계층에 대응하는 코딩 밴드를 알려주는 정보를 말한다. 스케일 밴드와 코딩 밴드는 실험에 의해 적절히 나누어지며 대응하는 스케일 팩터와 코딩 모델이 결정된다.

<40> 도 2는 본 발명의 바람직한 실시예에 따른 복호화 장치의 블럭도이다.

<41> 도 2를 참조하면, 복호화 장치는 네트워크 상황, 복호화 장치의 성능, 사용자 선택 등에 따라 결정된 타겟 계층까지 복호화함으로써 비트율을 조절할 수 있는 장치로서, 언패킹부(21), 역양자화부(22) 및 역변환부(23)를 포함한다.

<42> 언패킹부(21)는 비트스트림을 타겟 계층까지 언패킹하고 각 계층 별로 복호화한다. 즉, 각 계층 대응하는 스케일 팩터 정보, 코딩 모델 정보가 포함된 부가 정보를 복호화한 다음, 얻은 코딩 모델 정보를 기초로 각 계층에 속하는 부호화된 양자화 샘플들을 다시 복호화하여 양자화 샘플들을 얻는다. 본 실시예에서 언패킹부(21)는 스케일 팩터 정보와 코딩 모델 정보를 차분 복호화하고, 부호화된 양자화 샘플들은 허프만 복호화한다.

<43> 한편, 스케일 밴드 정보와 코딩 밴드 정보는 비트스트림의 헤더 정보로부터 얻거나, 각 계층 별 부가 정보를 복호화하여 얻을 수 있다. 대안적으로, 복호화 장치가 스케일 밴드 정보 및 코딩 밴드 정보를 미리 저장하고 있을 수도 있다. 역양자화부(22)는 각 계층의 양자화 샘플을 대응하는 스케일 팩터 정보에 따라 역양자화하여 복원한다. 역변환부(23)는 복원된 샘플을 주파수/시간 매핑하여 시간 영역의 PCM 오디오 데이터로 변환하여 출력한다. 본 실시예에서 역변환부(23)는 MDCT에 따른 역변환을 수행한다.

<44> 도 3은 본 발명에 따라 비트율을 조절할 수 있도록 계층 구조로 부호화된 비트스트림을 구성하는 프레임의 구조를 보여준다.

<45> 도 3을 참조하면, 본 발명에 따른 비트스트림의 프레임은 FGS(Fine Grain Scalability)를 위해 양자화 샘플과 부가 정보를 계층 구조에 맵핑시켜 부호화되어 있다. 즉, 하위 계층의 비트스트림이 상위 계층의 비트스트림에 포함되어 있는 계층 구조를 가진다. 각 계층에 필요 한 부가 정보들은 계층 별로 나뉘어서 부호화된다.

<46> 비트스트림의 선두에는 헤더 정보가 저장된 헤더 영역이 마련되고, 계층 0의 정보가 패킹되어 있으며, 상위 계층(enhancement layer)인 계층 1 내지 계층 N에 속하는 정보가 순서대로 패킹되어 있다. 헤더 영역에서부터 계층 0 정보까지를 기저 계층(base layer)이라고 부르고, 헤더 영역에서부터 계층 1 정보까지를 계층 1, 계층 2 정보까지를 계층 2라고 부른다. 마찬가지 방식으로, 최상위 계층은 헤더 영역에서부터 계층 N 정보까지, 즉 기저 계층에서부터 상위 계층인 계층 N까지를 말한다. 각 계층 정보로는 부가 정보와 부호화된 오디오 데이터가 저장되어 있다. 가령, 계층 2 정보로 부가 정보 2와 부호화된 양자화 샘플들이 저장되어 있다 . 여기서, N은 1 보다 크거나 같은 정수이다.

<47> 도 4는 부가 정보의 상세 구조를 보여준다.

<48> 도 4를 참조하면, 임의의 계층 정보로는 부가 정보와 부호화된 양자화 샘플들이 저장되어 있고, 본 실시예에서 부가 정보는 허프만 코딩 모델 정보, 양자화 팩터 정보, 채널에 대한 부가 정보와 기타 부가 정보를 포함한다. 허프만 코딩 모델 정보는 대응하는 계층에 속하는 양자화 샘플들의 부호화에 사용되거나 복호화에 사용되어야 할 허프만 코딩 모델에 대한 인덱스 정보를 말한다. 양자화 팩터 정보는 대응하는 계층에 속하는 오디오 데이터를 양자화하거나 역영자화하기 위한 양자화 스텝 사이즈를 알려준다. 채널에 대한 부가 정보란 M/S stereo 와 같은 채널에 대한 정보를 말한다. 기타 부가 정보는 M/S stereo의 채용 여부에 대한 플래그 정보 등을 말한다.

<49> 본 실시예에서, 비트 패킹부(14)는 허프만 코딩 모델 정보와 양자화 팩터 정보는 바로 전 밴드의 값의 차분값을 부호화하는 차분 부호화한다. 채널에 대한 부가 정보 등은 허프만 부호화한다.

<50> 도 5는 본 발명에 따른 부호화 방식을 개략적으로 설명하기 위한 참고도이다.

<51> 도 5를 참조하면, 부호화해야 할 양자화 샘플 전체가 3 개의 계층으로 구성되어 있다. 빛금친 사각형은 양자화 샘플들로 구성된 스펙트럼 라인을 나타내며, 실선은 스케일 밴드를 표시하며, 띠선은 코딩 밴드를 나타낸다. 계층 0에는 스케일 밴드 ①, ②, ③, ④ 및 ⑤가 속하며, 코딩 밴드 ①, ②, ③, ④ 및 ⑤가 속하고, 계층 1에는 스케일 밴드 ⑤ 및 ⑥이 속하며, 코딩 밴드 ⑥, ⑦, ⑧, ⑨ 및 ⑩이 속하고, 계층 2에는 스케일 밴드 ⑥ 및 ⑦이 속하며, 코딩 밴드 ⑪, ⑫, ⑬, ⑭ 및 ⑮가 속한다. 한편, 계층 0은 주파수 대역 ⑨까지 부호화하도록 고정되어 있고, 계층 1은 주파수 대역 ⑩까지 부호화하도록 고정되어 있으며, 계층 2는 주파수 대역 ⑮까지 부호화하도록 고정되어 있다.

<52> 먼저, 100비트 내에서 계층 0에 해당하는 양자화 샘플들을 해당하는 코딩 밴드 ①, ②, ③, ④ 및 ⑤에 정해져 있는 코딩 모델을 사용하여 부호화한다. 또한, 계층 0의 부가 정보로서 계층 0에 속하는 스케일 밴드 ①, ②, ③, ④, ⑤와 코딩 밴드 ①, ②, ③, ④, ⑤를 부호화한다. 계층 0의 샘플들을 심벌 단위로 부호화하면서 비트 수를 카운트하여 허용된 비트 범위, 즉 100비트를 넘어서면 계층 0의 부호화를 중단하고 계층 1을 부호화한다. 부호화되지 못한 계층 0의 샘플들은 계층 1 및 계층 2에 허용된 비트 범위에 여유가 생겼을 때 부호화한다.

<53> 다음으로, 계층 1에 속하는 코딩 밴드, 즉 코딩 밴드 ⑥, ⑦, ⑧, ⑨ 및 ⑩ 중 부호화하고자 하는 양자화 샘플이 속하는 코딩 밴드의 코딩 모델을 사용하여 계층 1에 속하는 양자화 샘플들을 부호화한다. 또한, 계층 1의 부가 정보로서 계층 1에 속하는 스케일 밴드 ⑤ 및 ⑥과 코딩 밴드 ⑥, ⑦, ⑧, ⑨ 및 ⑩을 부호화한다. 만일 계층 1에 해당하는 샘플들을 모두 부호화하고도 허용된 비트 범위, 즉 100비트가 되지 않을 경우에는 100비트가 다 챔 때까지 계층 0에서 부호화하지 못하였던 샘플을 부호화한다. 계층 1에 해당하는 샘플들을 심벌 단위로 부호화하면서 비트 수를 카운트하여 허용된 비트 범위, 즉 100비트를 넘어서면 계층 1의 부호화를 중단하고 계층 2의 부호화로 넘어간다.

<54> 마지막으로, 계층 2에 속하는 코딩 밴드, 즉 코딩 밴드 ⑪, ⑫, ⑬, ⑭ 및 ⑮ 중 부호화하고자 하는 양자화 샘플이 속하는 코딩 밴드의 코딩 모델을 사용하여 계층 2에 속하는 양자화 샘플을 부호화한다. 또한, 계층 2의 부가 정보로서 계층 2에 속하는 스케일 밴드 ⑥ 및 ⑦과 코딩 밴드 ⑪, ⑫, ⑬, ⑭ 및 ⑮를 부호화한다. 만일 계층 1에 해당하는 샘플들을 모두 부호화하고도 허용된 비트 범위, 즉 100비트가 되지 않을 경우에는 100비트가 다 챔 때까지 계층 0에서 부호화하지 못하였던 샘플을 부호화한다. 계층 1에 해당하는 샘플들을 심벌 단위로 부호화하면서 비트 수를 카운트하여 허용된 비트 범위, 즉 100비트를 넘어서면 계층 1의 부호화를 중단하고 계층 2의 부호화로 넘어간다.

화하면서 비트 수를 카운트하여 허용된 비트 범위, 즉 100비트를 넘어서면 계층 1의 부호화를 중단하고 계층 2의 부호화로 넘어간다.

<55> 만일 계층 0에서 허용된 비트 범위를 고려하지 않고 해당하는 양자화 샘플을 모두 부호화해 버린다면, 다시 말해 부호화된 비트 수가 이미 허용된 비트 범위, 즉 100비트를 초과하였는데도 불구하고 모두 부호화한다면 결국 다음 계층인 계층 1에 허용된 비트 범위의 적어도 일부를 차용하는 셈이 되어 정작 계층 1에 속하는 양자화 샘플들을 부호화할 수 없게 되는 일이 발생된다. 따라서, 비트율 조절 가능(scalable)하게 복호화할 경우 계층 1까지만 복호화한다면 계층 1의 주파수 ⑥까지 부호화되지 못했기 때문에 복호화된 양자화 샘플들은 주파수 ⑥ 이하에서 오르락 내리락하는 모습을 띄게 된다. 이 때 음질이 열화되는 버디 효과(birdy effect)가 나타난다.

<56> 한편, 복수개의 계층(타겟 계층)을 결정할 때 부호화해야 할 오디오 데이터 전체의 크기를 고려하여 비트 범위가 할당되므로 전체적으로 부호화해야 할 비트 범위가 모자라서 부호화하지 못하는 경우는 발생되지 않는다.

<57> 복호화 과정 또한 부호화와 마찬가지로 그 역과정을 수행하면서 허용하는 비트 범위에 따라 비트 수를 카운트하기 때문에 계층 1로 복호화할 시점을 알아낼 수 있다.

<58> 도 6은 본 발명에 따른 부호화 방식을 보다 구체적으로 설명하기 위한 참고도이다.

<59> 본 발명에 따르면 비트패킹부(14)는 각 계층에 해당하는 양자화된 샘플들을 비트 플래인(bit-plane) 부호화와 헤프만 부호화를 채용하여 부호화한다.

<60> 복수개의 양자화 샘플들을 비트 플래인 상에 매핑시켜 이진 데이터로 나타내고 이진 데이터의 최상위비트(Most Significant Bit, MSB)들로 구성된 심볼부터 최하위 비트(Least

Significant Bit, LSB)들로 구성된 심볼의 순서로 해당 계층에 할당된 비트 범위 내에서 부호화한다. 비트 플래인 상에서 중요한 정보는 먼저 부호화하고 상대적으로 덜 중요한 정보는 나중에 부호화할 수 있다. 부호화 과정에서 각 계층에 해당하는 비트율과 계층별 주파수 대역을 고정하여 버디 효과라고 부르는 왜곡(distortion)을 줄이기 위함이다.

<61> 도 6은 msb들로 구성된 심벌의 비트수가 4 이하일 경우에 대한 부호화의 예로서, 양자화 샘플들 9, 2, 4, 0을 비트 플래인에 매핑하면 각각 1001b, 0010b, 0100b, 0000b의 이진 데이터로 표시된다. 즉, 본 실시예에서 비트 플레인 상에서 부호화 단위가 되는 부호화 블럭의 크기는 4*4이다.

<62> 최상위비트들 msb로 구성된 심볼은 "1000b"이고, 그 다음 비트들 msb-1로 구성된 심볼은 "0010b"이며, 그 다음 비트들 msb-2로 구성된 심볼은 "0100b"이고, 최하위비트들 msb-3로 구성된 심볼은 "1000b"이다.

<63> 허프만 부호화를 위한 허프만 모델 정보, 즉 코드북 인덱스는 [표 1]과 같다.

<64>

【표 1】

부가 정보	중요도(significance)	히프만 모델
0	0	0
1	1	1
2	1	2
3	2	3
		4
4	2	5
		6
5	3	7
		8
		9
6	3	10
		11
		12
7	4	13
		14
		15
		16
8	4	17
		18
		19
		20
9	5	*
10	6	*
11	7	*
12	8	*
13	9	*
14	10	*
15	11	*
16	12	*
17	13	*
18	14	*
*	*	*

<65> [표 1]에 따르면, 같은 중요도(significance)(본 실시예에서는 msb)에 대해서도 두 개의 모델이 존재하는 것을 볼 수 있는데 이는 서로 다른 분포를 보이는 양자화 샘플들에 대해 두 개의 모델을 생성했기 때문이다.

<66> 도 6의 예를 [표 1]에 따라 부호화하는 과정을 보다 구체적으로 설명하면 다음과 같다.

<67> 심벌의 비트수가 4 이하일 경우 본 발명에 따른 허프만 부호화는 【수학식 1】에 따른다.

<68> 【수학식 1】 허프만 코드값

<69> = HuffmanCodebook[코드북 인덱스][상위 비트 플레인][심볼]

<70> 즉, 허프만 부호화는 3 개의 입력변수로서, 코드북 인덱스, 상위 비트 플레인 및 심볼을 가진다. 코드북 인덱스는 【표 1】로부터 얻어진 값을 말하며, 상위 비트 플레인은 비트 플레인 상에서 현재 부호화하고자 하는 심볼의 바로 위의 심벌을 가리킨다. 심볼은 현재 부호화하고자 하는 심볼을 말한다.

<71> 도 6의 예에서 허프만 모델은 msb가 4이므로 13-16 또는 17-20이 선택된다. 부호화될 부가 정보가 7이라면,

<72> msb들로 구성된 심벌의 코드북 인덱스는 16,

<73> msb-1들로 구성된 심벌의 코드북 인덱스는 15,

<74> msb-2들로 구성된 심벌의 코드북 인덱스는 14,

<75> msb-3들로 구성된 심벌의 코드북 인덱스는 13

<76> 이 된다.

<77> 한편, msb들로 구성된 심벌은 상위 비트 플레인의 데이터를 가지고 있지 않으므로 상위 비트플레인 값을 0이라 가정하면, HuffmanCodebook[16][0b][1000b]의 코드로 부호화된다.

msb-1들로 구성된 심벌은 상위 비트 플레인이 1000b이므로 HuffmanCodebook[15][1000b][0010b]

의 코드로 부호화된다. msb-2들로 구성된 심벌은 상위 비트 플레인이 0010b이므로

HuffmanCodebook[14][0010b][0100b]의 코드로 부호화된다. msb-3들로 구성된 심벌은 상위 비트 플레인이 0100b이므로 HuffmanCodebook[13][0100b][1000b]의 코드로 부호화된다.

<78> 비트 패킹부(14)는 심벌 단위로 부호화한 다음 부호화된 총 비트수를 카운트하고 사용 가능한 비트수와 비교하여 부호화된 비트수가 해당 계층에서 사용가능한 비트수를 초과할 경우 부호화를 중지한다. 부호화되지 못하고 남겨진 비트는 다음 계층에 여유 공간이 생길 때 부호화하여 집어넣는다. 해당 계층에 할당된 양자화 샘플들을 모두 부호화하고 나서도 사용가능한 비트수가 남을 경우, 즉 여유 공간이 생길 경우에는 하위 계층에서 부호화되지 못하고 남은 양자화 샘플을 부호화한다.

<79> 한편, msb들로 구성된 심벌의 비트수가 5 이상일 경우에는 현재 비트 플래인 상의 위치를 이용해 허프만 코드값을 결정한다. 즉, 중요도가 5 이상일 경우에는 각각의 비트 플래인 상의 데이터는 통계적으로 큰 차이를 보이지 않기 때문에 모두 동일한 허프만 모델을 사용하여 허프만 부호화한다. 즉, 비트 플래인 당 하나의 허프만 모델이 존재한다.

<80> 중요도가 5 이상일 경우(심벌의 비트수가 5 이상일 경우) 본 발명에 따른 허프만 부호화는 [수학식 2]에 따른다.

<81> 【수학식 2】 허프만 코드값 = $20 + bpl$

<82> 여기서, bpl 은 현재 코딩하고자 하는 비트 플래인의 인덱스를 가리키며, 따라서 1 이상의 정수값을 가진다. 20은 [표 1]의 부가 정보 8에 대응하는 허프만 모델의 마지막 인덱스가 20이므로 인덱스가 21부터 시작되도록 하기 위해 더해주는 값이다. 따라서 코딩 밴드에 대한 부가 정보는 단순히 중요도만을 나타낸다. 아래의 [표 2]에서 허프만 모델은 현재 부호화하고자 하는 비트 플래인의 인덱스에 따라 결정된다.

<83>

【표 2】

부가 정보	중요도(significance)	허프만 모델
9	5	21-25
10	6	21-26
11	7	21-27
12	8	21-28
13	9	21-29
14	10	21-30
15	11	21-31
16	12	21-32
17	13	21-33
18	14	21-34
19	15	21-35

<84> 한편, 부가 정보 중 양자화 팩터 정보와 허프만 모델 정보는 대응하는 코딩 밴드에 대해 차분 부호화(DPCM)를 수행한다. 양자화 팩터 정보를 부호화할 때 차분 부호화의 초기값은 프레임의 헤더 정보에 8 bit로 표현된다. 허프만 모델 정보에 대한 차분 부호화의 초기값은 0으로 세팅한다.

<85> 본 발명에 따른 부호화 방식과 종래의 BSAC와의 차이점은 다음과 같다. 첫째, BSAC는 비트 단위로 부호화하였으나 본 발명에 따르면 심볼 단위로 부호화한다. 둘째, BSAC는 산술 부호화하였으나 본 발명은 허프만 부호화한다. 산술 부호화는 압축 계인은 높지만 많은 계산량이 요구되어 complexity와 비용이 높아지는 문제점이 있다. 따라서, 본 발명에서는 비트 단위로 부호화하지 않고 심벌 단위로 허프만 부호화함으로써 complexity와 비용을 낮추었다.

<86> 비트율을 조정하기 위해서는, 즉 scalability를 적용할 경우 한 프레임에 해당하는 비트스트림을 각 계층에서 사용가능한 비트수를 고려하여 잘라줌으로써 적은 데이터만으로도 복호화할 수 있게 된다. 예를 들어, 최고 비트율이 96kbps이고 대응하는 비트스트림의 크기가

2096비트일 경우, 48kbps에 해당하는 비트스트림만을 복호화하고자 할 때에는 비트스트림을 1048 비트만 취함으로써 48kbps에 해당하는 복호화된 오디오 데이터를 얻을 수 있다.

<87> 상기한 구성을 기초로 본 발명에 따른 부호화 방법 및 복호화 방법을 설명하면 다음과 같다.

<88> 부호화 장치는 PCM 오디오 데이터를 읽어들여 메모리(도시되지 않음)에 저장하고, 저장된 PCM 오디오 데이터를 이용하여 심리음향모델을 거쳐 마스킹 문턱치와 부가 정보를 구한다. PCM 오디오 데이터는 시간 영역의 신호이므로 이를 주파수 영역의 신호로 MDCT 변환한다. 다음으로, 부호화 장치는 변환된 신호를 양자화 밴드 정보와 양자화 팩터 정보에 따라 양자화하여 양자화 샘플을 얻는다. 양자화 샘플을 전술한 바와 같이 비트 분할 부호화, 심벌 단위 부호화 및 헤프만 부호화를 채용하여 부호화하여 패킹한다.

<89> 도 7은 본 발명의 바람직한 실시예에 따른 부호화 방법을 설명하기 위한 플로우챠트이다.

<90> 도 7을 참조하면, 부호화 장치의 비트 패킹부(14)에 의해 양자화 샘플을 부호화하여 패킹하는 과정은 다음과 같다.

<91> 먼저, 비트 패킹부(14)는 주어진 타겟 비트율과 부가 정보를 기초로 각 계층에 해당하는 정보를 추출해낸다. 이 과정은 단계 701 내지 703을 통해 수행된다. 즉, 각 계층에 분할하는 기준이 되는 cut-off 주파수를 구하고(701단계), 각 계층에 해당하는 양자화 밴드 정보와 코딩 밴드 정보를 구하여(702단계), 부호화되어야 할 비트를 이용하여 계층 당 부호화 가능한 비트 범위를 할당한다(703단계).

- <92> 다음으로, 기반 계층에 해당하는(704단계) 양자화 밴드 정보와 코딩 밴드 정보를 포함하여 부가 정보를 부호화한다(705단계).
- <93> 다음으로, 기반 계층에 해당하는 양자화 샘플을 비트 플래인 상에 매핑하고 4*4 블럭 단위로 msb들로 구성된 심벌부터 부호화하면서(706단계), 부호화된 비트수를 카운트하여 카운트된 비트수가 해당 계층에서 사용가능한 비트 범위를 초과하면(707단계), 현재 계층의 부호화를 중지하고 다음 계층으로 넘어간다. 카운트된 비트수가 해당 계층에서 사용가능한 비트 범위 이하이면(707단계) 다음 계층에 대해 상기 705단계로 넘어간다(709단계). 708단계는 기반 계층의 경우에는 그 보다 낮은 하위 계층이 없으므로 수행되지 않으나 이후 계층에 대해서는 수행된다. 상기 과정을 거쳐서 타겟 계층까지 모두 부호화한다.
- <94> 상기 706단계, 즉 양자화 샘플을 부호화하는 과정은 다음과 같다.
- <95> 1. 각 계층에 해당하는 양자화 샘플을 N 개의 단위로 묶어 비트 플래인 상에 매핑한다.
- <96> 2. 매핑된 이진 데이터들의 msb들로 구성된 심볼부터 차례대로 허프만 부호화한다.
- <97> 2의 과정을 다시 세분하여 설명하면 다음과 같다.
- <98> 2.1. 현재 부호화하고자 하는 심볼에 대응하는 하나의 스칼라값(curVal)을 구한다.
- <99> 2.2. 상위 비트 플래인, 즉 현재 부호화하고자 하는 심볼의 윗단에 존재하는 심볼에 대응하는 스칼라값(upperVal)에 대응하는 허프만 코드를 구하여 구해진 코드로 부호화한다.
- <100> 부가 정보 중 양자화 팩터 정보와 허프만 모델 정보는 대응하는 코딩 밴드에 대해 차분 부호화(DPCM)를 수행한다. 양자화 팩터 정보를 부호화할 때 차분 부호화의 초기값은 프레임의 헤더 정보에 8 bit로 표현된다. 허프만 모델 정보에 대한 차분 부호화의 초기값은 0으로 세팅 한다.

- <101> 도 8은 본 발명의 바람직한 실시예에 따른 복호화 방법을 설명하기 위한 플로우챠트이다.
- <102> 도 8을 참조하면, 복호화 장치는 계층 구조로 부호화된 오디오 데이터로 구성된 비트스트림을 수신하여 프레임 별로 마련된 헤더 정보를 복호화한다. 다음으로, 첫번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 복호화한 다음(801단계), 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 복호화하여 양자화 샘플을 얻는다(802단계). 얻어진 양자화 샘플을 상기 스케일 팩터 정보를 참조하여 역양자화하고(803단계), 역양자화된 샘플을 역변환한다(804단계). 미리 결정된 타겟 계층에 대한 복호화가 완료될 때까지 매 회 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 801단계 내지 804단계를 반복 수행한다.
- <103> 도 9는 본 발명의 바람직한 실시예에 따른 복호화 방법을 설명하기 위한 플로우챠트이다.
- <104> 도 9를 참조하면, 복호화 장치는 계층 구조로 부호화된 오디오 데이터로 구성된 비트스트림을 수신하여 프레임 별로 마련된 헤더 정보로부터 각 계층에 해당하는 컷-오프 주파수를 복호화하고(901단계), 헤더 정보로부터 각 계층에 해당하는 양자화 밴드 정보와 코딩 밴드 정보를 복호화하여 알아낸 다음(902단계), 각 계층별로 사용가능한 비트 범위를 알아낸다(903단계). 이어서, 기반 계층에 대한(904단계) 부가 정보를 복호화하고(905단계), 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위만큼 심벌 단위로 복호화하여 양자화된 샘플을 얻는다(906단계). 미리 결정된 타겟 계층까지(907단계), 계층을 하나씩 증가시키면서 상기 905단계 내지 906단계를 반복한다. 상기 901단계 내지 903단계는 컷-오프 주파수, 양자화 밴드 정보, 코딩 밴드 정보 및 비트 범위는 수신된

비트스트림의 프레임마다 저장된 헤더 정보로부터 얻는 대신 복호화 장치가 미리 가지고 있을 수도 있고, 그와 같은 경우 복호화 장치는 저장된 정보를 읽어들여서 각 정보를 얻는다.

【발명의 효과】

- <105> 전술한 바와 같이, 본 발명에 따르면 비트 분할한 다음 심벌 단위로 부호화함으로써 top-down방식으로 비트율을 조절할 수 있는 scalability를 제공하므로 부호화 장치의 계산량이 scalability를 제공하지 않는 장치의 그것에 비해 크게 많지 않다. 즉, 본 발명에 따르면 FGS(Fine Grain Scalability)를 제공하면서도 complexity가 보다 낮은, 비트율 조절가능한 오디오 부호화 방법, 복호화 방법, 그 부호화 장치 및 복호화 장치를 얻을 수 있다.
- <106> 나아가, 산술 부호화를 이용하는 MPEG-4 Audio BSAC에 비해 허프만 부호화를 이용하는 본 발명의 부호화 장치/복호화 장치는 비트 패킹/언패킹(bit packing/unpacking) 과정에서 계산량이 크게 줄어든다. FGS를 제공하기 위해 본 발명에 따른 비트 패킹을 수행하여도 오버헤드가 적어 부호화 이득 측면에서 scalability를 제공하지 않은 경우와 거의 유사하다.
- <107> 또한, 본 발명에 따른 장치는 계층 구조를 가지므로 서버 측에서 비트율 조절가능하도록 비트스트림을 재생성하는 과정이 매우 단순하며 따라서 변환 부호화 장치의 복잡도가 낮고, 네트워크를 통한 오디오스트림 전송시 사용자의 의지 혹은 네트워크 환경에 따라 전송 비트율을 변경하여 전송함으로써 끊김없는 서비스의 제공이 가능하다.
- <108> 나아가, 용량의 제한을 갖는 정보저장매체에 저장할 때 파일 사이즈를 임의로 조절하여 저장할 수 있게 된다. 비트율이 낮아지면 대역이 제한되어 있기 때문에 주로 부호화/복호화 장치의 복잡성의 대부분을 차지하는 필터의 복잡성이 상당히 감소하기 때문에 비트율에 반비례해서 부호화 장치/복호화 장치의 실제 복잡성도 감소하게 된다.

【특허청구범위】**【청구항 1】**

- (a) 복수개의 계층에 상응하도록 오디오 데이터를 분할하는 단계;
- (b) 상기 복수개의 계층에 각각 상응하는 스케일 밴드 정보 및 코딩 밴드 정보를 얻는 단계;
- (c) 첫번째 계층에 상응하는 스케일 밴드 정보 및 코딩 밴드 정보에 따른 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 부호화하는 단계;
- (d) 상기 첫번째 계층에 해당하는 오디오 데이터를 대응하는 스케일 팩터 정보를 참조하여 양자화하여 양자화 샘플을 얻는 단계;
- (e) 얻어진 복수개의 양자화 샘플을 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 허프만 부호화하는 단계; 및
- (f) 상기 복수개의 계층에 대한 부호화가 완료될 때까지 매 회 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 (c)단계 내지 (f)단계를 반복하여 계층 구조를 갖도록 패킹하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

【청구항 2】

제1항에 있어서,

상기 (f)단계 이전에

(e) 상기 복수개의 계층에 대해 각각 사용가능한 비트 범위를 구하는 단계를 포함하고,

상기 (f)단계는

부호화된 비트 수를 카운트하여, 카운트된 비트 수가 대응하는 비트 범위를 초과하면 부호화를 중지하며, 할당된 양자화 샘플을 전부 부호화한 이후에도 카운트된 비트 수가 대응하는 비트 범위보다 작으면 하위 계층에서 부호화하지 못하고 남은 비트를 상기 비트 범위가 허용하는 만큼 부호화하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

【청구항 3】

제1항 또는 제2항에 있어서,

상기 (a)단계는

(a1) 오디오 데이터를 MDCT 변환하는 단계; 및

(a2) MDCT 변환된 오디오 데이터를 컷-오프 주파수를 참조하여 상기 복수개의 계층에 상응하도록 분할하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

【청구항 4】

제1항 또는 제2항에 있어서,

상기 (f)단계는

(f1) 복수개의 양자화 샘플을 비트 플래인 상에 매핑하는 단계; 및

(f2) 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위만큼 심벌 단위로 부호화하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

【청구항 5】

제4항에 있어서,

상기 (f1)단계는

N 개의 양자화 샘플을 비트 플래인 상에 매핑하는 단계이고,

상기 (f2)단계는

N 비트의 이진 데이터로 구성된 상기 심벌에 대응하는 하나의 스칼라 값을 구하는 단계;

및

상기 N 비트의 이진 데이터, 구해진 스칼라 값과 상기 비트 플래인 상에서 현재의 심벌에 대한 상위 심벌에 대응하는 스칼라 값을 참조하여 헤프만 부호화를 수행하는 단계를 포함하는 것을 특징으로 하는 부호화 방법(N은 정수).

【청구항 6】

계층 구조로 부호화된 오디오 데이터를 비트율 조절이 가능하도록 복호화하는 방법에 있어서,

- (a) 첫번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 복호화하는 단계;
- (b) 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 헤프만 복호화하여 양자화 샘플을 얻는 단계;
- (c) 얻어진 양자화 샘플을 상기 스케일 팩터 정보를 참조하여 역양자화하는 단계;
- (d) 역양자화된 샘플을 MDCT 역변환하는 단계; 및
- (e) 미리 결정된 계층까지 복호화가 완료될 때까지 매 회 상기 계층 앞에 부가된 서수를 1 씩 증가시키면서 상기 (a)단계 내지 (d)단계를 반복 수행하는 단계를 포함하는 것을 특징으로 하는 복호화 방법.

【청구항 7】

제6항에 있어서,

상기 (b)단계는

(b1) 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위 만큼 심벌 단위로 복호화하는 단계; 및
(b2) 복호화된 심벌이 배열된 비트 플래인으로부터 양자화된 샘플을 얻는 단계를 포함하는 것을 특징으로 하는 복호화 방법.

【청구항 8】

제7항에 있어서,

상기 (b1)단계는

복호화된 심벌로 구성된 $4*N$ 비트 플래인을 얻는 단계이고,

상기 (b2)단계는

상기 $4*N$ 비트 플래인으로부터 N 개의 양자화 샘플을 얻는 단계를 포함하는 것을 특징으로 하는 복호화 방법(N 은 정수).

【청구항 9】

계층 구조로 부호화된 오디오 데이터를 비트율 조절이 가능하도록 복호화하는 장치에 있어서,

첫번째 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 복호화하고, 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하

위 비트들로 구성된 심벌의 순서로 심벌 단위로 허프만 복호화하여 양자화 샘플을 얻는 언패킹

부:

얻어진 양자화 샘플을 상기 스케일 팩터 정보를 참조하여 역양자화하는 역양자화부; 및
역양자화된 샘플을 역변환하는 역변환부를 포함하는 것을 특징으로 하는 복호화 장치.

【청구항 10】

제9항에 있어서,

상기 언패킹부는

최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층
이 허용하는 비트 범위 만큼 심벌 단위로 복호화하고, 복호화된 심벌이 배열된 비트 플래인으
로부터 양자화된 샘플을 얻는 것을 특징으로 하는 복호화 장치.

【청구항 11】

제10항에 있어서,

상기 언패킹부는

복호화된 심벌로 구성된 $4*N$ 비트 플래인을 얻은 다음 상기 $4*N$ 비트 플래인으로부터 N
개의 양자화 샘플을 얻는 것을 특징으로 하는 복호화 장치(N은 정수).

【청구항 12】

오디오 데이터를 비트율 조절이 가능하도록 부호화하는 장치에 있어서,

상기 오디오 데이터를 MDCT 변환하는 변환부;

각 계층에 해당하는, MDCT 변환된 오디오 데이터를 상기 스케일 팩터 정보를 참조하여
양자화하여 양자화 샘플을 출력하는 양자화부; 및

각 계층에 상응하는 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 차분 부호화하고, 상기 양자화부로부터의 복수개의 양자화 샘플을 상기 코딩 모델 정보를 참조하여 최상위 비트들로 구성된 심벌에서부터 최하위 비트들로 구성된 심벌의 순서로 심벌 단위로 허프만 부호화하는 패킹부를 포함하는 것을 특징으로 하는 부호화 장치.

【청구항 13】

제12항에 있어서,

상기 패킹부는 복수개의 계층에 각각 상응하는 스케일 뱀드 정보 및 코딩 뱀드 정보를 얻어 각 계층에 상응하는 스케일 뱀드 정보 및 코딩 뱀드 정보에 따른 스케일 팩터 정보 및 코딩 모델 정보를 포함하는 부가 정보를 부호화하는 것을 특징으로 하는 부호화 장치.

【청구항 14】

제12항에 있어서,

상기 패킹부는

부호화된 비트 수를 카운트하여, 카운트된 비트 수가 각 계층에 허용되는 비트 범위를 초과하면 부호화를 중지하며, 할당된 양자화 샘플을 전부 부호화한 이후에도 카운트된 비트 수가 대응하는 비트 범위보다 작으면 하위 계층에서 부호화하지 못하고 남은 비트를 상기 비트 범위가 허용하는 만큼 부호화하는 것을 특징으로 하는 부호화 장치.

【청구항 15】

제12항 내지 제13항 중 어느 한 항에 있어서,

상기 패킹부는 MDCT 변환된 오디오 데이터를 상기 컷-오프 주파수를 참조하여 복수개의 계층에 상응하도록 분할하는 것을 특징으로 하는 부호화 장치.

【청구항 16】

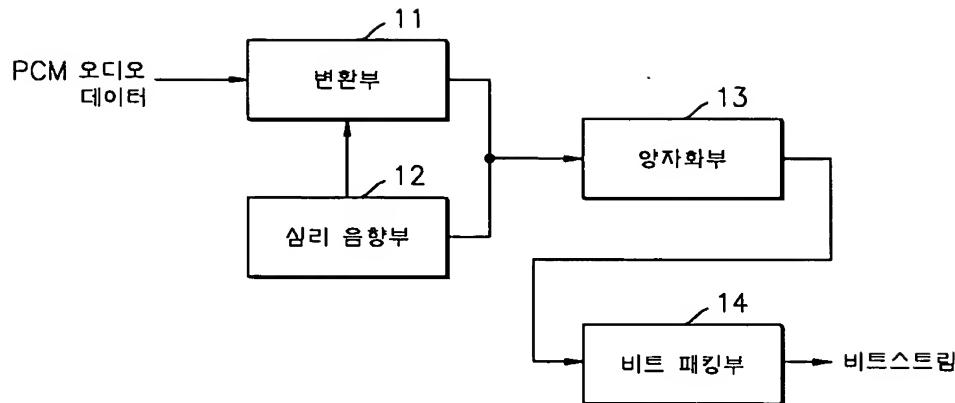
제12항 내지 제13항 중 어느 한 항에 있어서,
상기 패킹부는 복수개의 양자화 샘플을 비트 플래인 상에 매핑하고, 최상위 비트들로 구성된 심벌부터 최하위 비트들로 구성된 심벌의 순서로 대응하는 계층이 허용하는 비트 범위만큼 심벌 단위로 부호화하는 것을 특징으로 하는 부호화 장치.

【청구항 17】

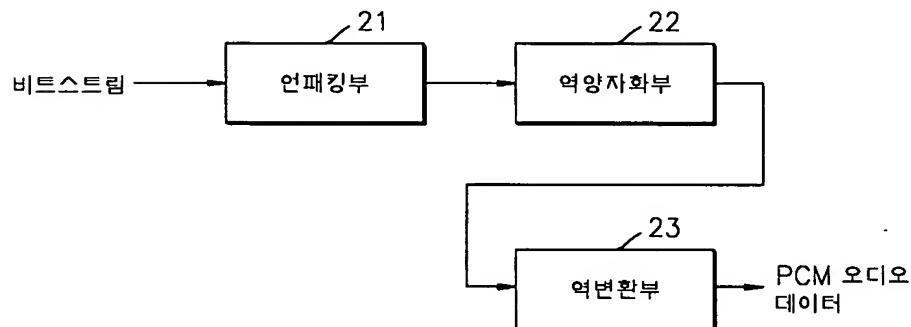
제16항에 있어서,
상기 패킹부는 K 개의 양자화 샘플을 비트 플래인 상에 매핑하고, K 비트의 이진 데이터로 구성된 상기 심벌에 대응하는 하나의 스칼라 값을 구한 다음 상기 K 비트의 이진 데이터, 구해진 스칼라 값과 상기 비트 플래인 상에서 현재의 심벌에 대한 상위 심벌에 대응하는 스칼라 값을 참조하여 허프만 부호화하는 것을 특징으로 하는 부호화 장치(K는 정수).

【도면】

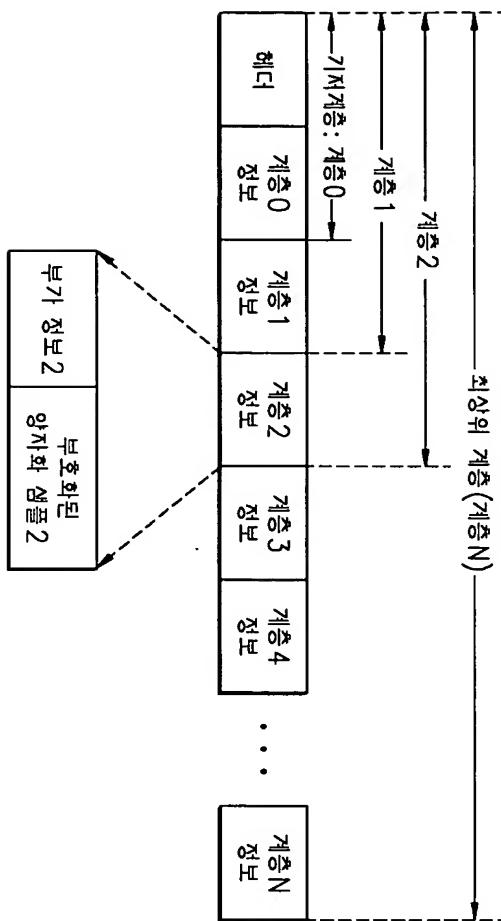
【도 1】



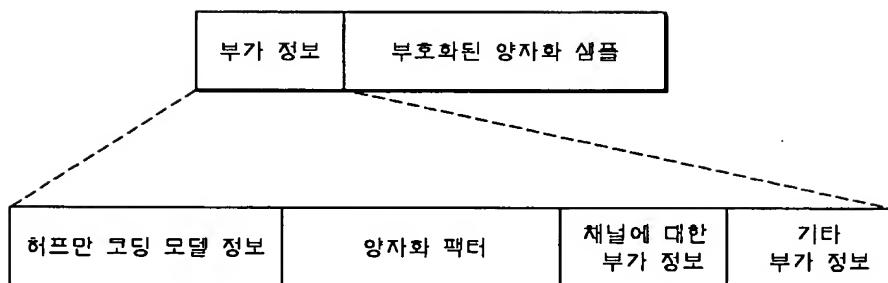
【도 2】



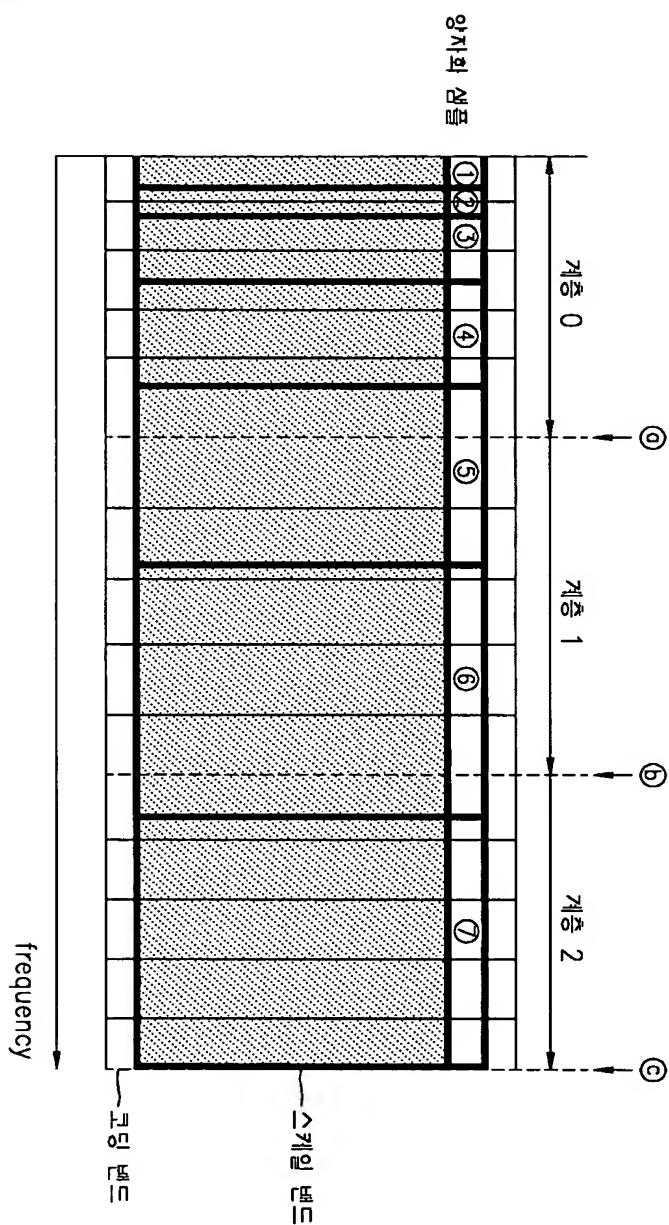
【도 3】



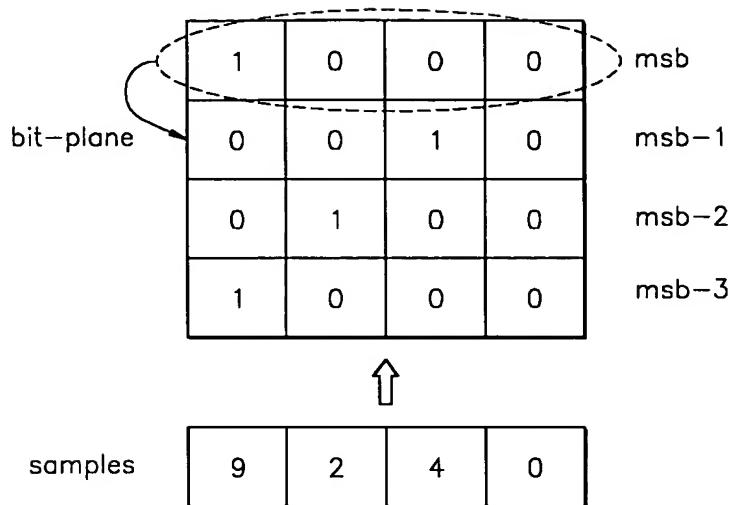
【도 4】



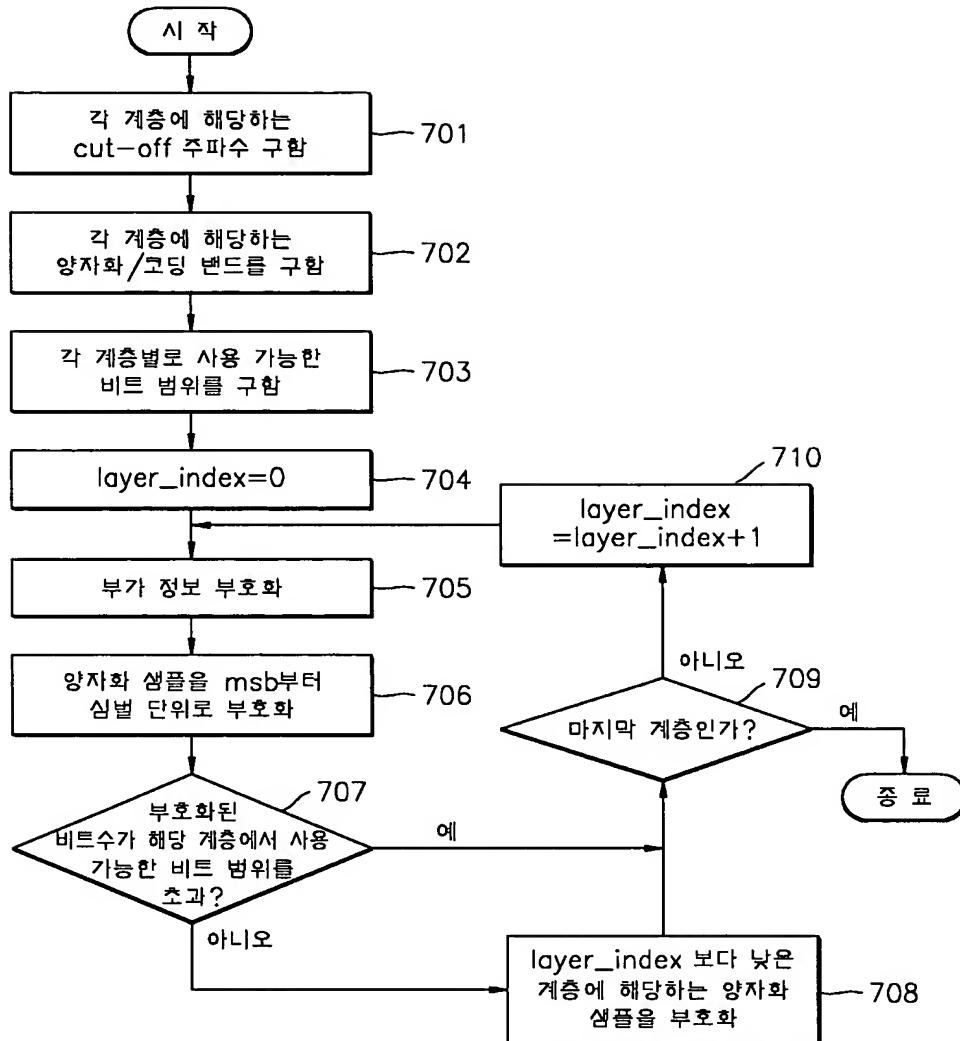
【도 5】



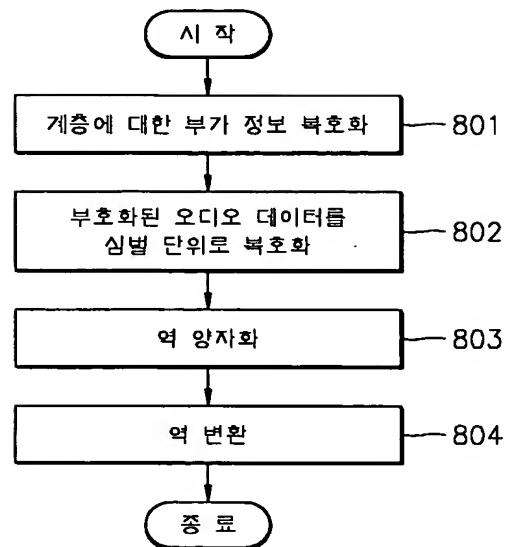
【도 6】



【도 7】



【도 8】



【도 9】

